SCIENCESPRINGDAY

FCT FACULDADE DE CIÊNCIAS E TECNOLOGIA UNIVERSIDADE NOVA DE LISBOA

Departamento de Informática

Provenance in the Semantic Web

Knowledge and Information Systems Group

CENTRIA

Carlos Damásio

(Associate Professor)

Degrees:
1996: PhD in Computer Science from UNL

Objectives

Companies, researchers, governmental organizations, and users are publishing huge sets of structured data in the form of Linked Open Data in the Semantic Web. This data can be introduced by humans or generated by automatic processes that gather, link and combine together the several pieces of information and knowledge. Knowing the sources is fundamental!

The objective is to develop theories and tools that can inform users about data provenance, namely the sources used and how the the information has been obtained or constructed. The focus is on data generated by rules specified by the standard languages and models of W3C.

Semantic Web

Methodology

Linked Open Data is published in the form of Resource Description Framework (RDF) graphs, which can be constructed with the SPARQL querying language for RDF or with the Rule Interchange Format (RIF). Provenance for reasoning with the absence of information is the challenge.

- Study relevant provenance models for relational databases
- Extend the models to deal with the non-monotonic constructs of SPARQL
- Generalise to the RIF language, and provide a declarative semantics
- Implement the data provenance model for RIF
- Apply and assess the model in large datasets

Expected Results

A fully usable data provenance model, semantics and tools for the querying languages of the Semantic Web, including:

- An abstract algebraic model capable of capturing provenance for all the constructs of the SPARQL and RIF languages.
- A data provenance semantics for knowledge bases specified in SPARQL and RIF.
- An extended SPARQL engine for storing and producing annotations according to the previous abstract model, and publish them in the form of Linked Open Data using the PROV-O ontology.

RDF

Funding:

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA

ERRO. efficient reasoning with rules and ontologies